

云斑车蝗线粒体基因组全序列测定与分析

党江鹏, 刘 念, 叶 伟, 黄 原*

(陕西师范大学生命科学学院, 西安 710062)

摘要: 采用长距 PCR 扩增及保守引物步移法并结合克隆测序测定并注释了云斑车蝗 *Gastrimargus marmoratus* (Thunberg) 的线粒体基因组全序列。结果表明: 云斑车蝗线粒体基因组全序列为 15 904 bp (GenBank 登录号为 EU527334), A + T 含量略高于非洲飞蝗 *Locusta migratoria*, 为 76.04%, 包括 13 个蛋白质编码基因, 22 个 tRNA 基因, 2 个 rRNA 基因和一段 1 057 bp 的 A + T 富集区。蛋白质基因的起始密码子中, 除 CO I 和 ND5 为 TTG 以外, 均为昆虫典型的起始密码子 ATN。ND5 基因使用了不完全终止密码子 T, 其余基因均为典型的 TAA 或 TAG。预测了 22 个 tRNA 基因的二级结构, 发现 tRNA^{Ser}(AGN) 缺少 DHU 臂, tRNA^{Ser}(UGY) 的反密码子环上有 9 个碱基。预测了云斑车蝗 12S 和 16S rRNA 二级结构, 分别包括 3 个结构域 30 个茎环和 6 个结构域 44 个茎环。A + T 富集区含有 3 个串联重复序列。

关键词: 直翅目; 蝗总科; 云斑车蝗; 线粒体基因组; RNA 二级结构; 重复序列

中图分类号: Q966 **文献标识码:** A **文章编号:** 0454-6296(2008)07-0671-10

Complete mitochondrial genome sequence of *Gastrimargus marmoratus* (Thunberg) (Orthoptera: Acridoidea)

DANG Jiang-Peng, LIU Nian, YE Wei, HUANG Yuan* (School of Life Sciences, Shaanxi Normal University, Xi'an 710062, China)

Abstract: The complete sequence of *Gastrimargus marmoratus* (Thunberg) mitochondrial genome was determined using long PCR and conserved primers walking approaches. The results showed that the entire mitochondrial genome of *G. marmoratus* (GenBank accession no. EU527334) is 15 904 bp in length with 76.5% A + T content. It contained 13 protein-coding genes, 2 ribosomal RNA genes, 22 tRNA genes, and an A + T-rich region of 1 057 bp. All protein-coding genes of the *G. marmoratus* mitochondrial genome start with a typical ATN codon, except for CO I and ND5 with TTG. The usual TAA and TAG termination codons were found for twelve protein-coding genes. However, the ND5 genes have an incomplete termination codon (T). With respect to the secondary structures of 22 tRNA genes, the absence of DHU arm of tRNA^{Ser}(AGN) was observed as in *Locusta migratoria*. tRNA^{Ser}(UGY) with 9 bases in anticodon loop was found, compared with typical 7 bases in others. The predicted secondary structures of 12S rRNA and 16S rRNA include 30 helices in 3 structural domains and 44 helices in 6 domains, respectively. The A + T content of the A + T-rich region is 76.5%, relatively higher than that of *L. migratoria*; three tandemly repeated DNA sequences were found in the A + T-rich region in *G. marmoratus*.

Key words: Orthoptera; Acridoidea; *Gastrimargus marmoratus*; mitochondrial genome; RNA secondary structure; tandemly repeated sequences

线粒体是大多数真核细胞进行有氧呼吸的细胞器, 拥有自身的遗传物质即线粒体基因组。典型的后生动物线粒体基因组编码 37 个基因, 包括 13 个

蛋白质基因, 22 个 tRNA 基因, 2 个 rRNA 基因和一段 A + T 富集区 (Wolstenholme, 1992)。线粒体基因组的蛋白质编码基因与核基因组相比进化速率较

基金项目: 国家自然科学基金项目 (30470238, 30670279)

作者简介: 党江鹏, 男, 1982 年生, 硕士研究生, 研究方向为昆虫分子系统学, E-mail: djiangpeng@stu.snnu.edu.cn

* 通讯作者 Author for correspondence, E-mail: yuanyh@snnu.edu.cn

收稿日期 Received: 2008-01-02; 接受日期 Accepted: 2008-05-12

快,因而被广泛用于科、属、种或种群水平的系统发育研究,比如由于蛋白质基因密码子第三位点具有极高的替换速率,被用于种和种群的研究(Navajas *et al.*, 1996)。与蛋白质基因相比,线粒体 rRNA 基因较为保守,尤其是 12S rRNA 的高度保守性可用于高级阶元的进化关系研究(Ballard *et al.*, 1992); 16S rRNA 的进化速率比 12S 稍快,用于中级分类阶元的研究(Black and Piesman, 1994; Kambhampati *et al.*, 1996)。线粒体 DNA 中的 A + T 富集区是高变区,该区域甚至在种内的不同个体之间也可能存在变化,因而主要用于种、亚种或种群的研究(Zhang and Hewitt, 1997)。

然而,将 rRNA 作为分子标记应用于系统发育和进化研究上时,面临着如何正确比对序列和选择进化模型两方面的困难。RNA 二级结构主要由茎区(stem regions)和环区(loop regions)组成,由于环区发生的插入/缺失(indels)比较多,序列的长度变化较大,所以编码 RNA 基因序列较编码蛋白质基因序列更难于正确比对。另一方面,茎区的序列是配对的,其碱基替换受保守二级结构的限制,在进化上表现出补偿性替换型式。因此,普通的 DNA 进化模型不能很好地拟合 RNA 序列的进化,只有得到二级结构方面的信息,将二级结构保守的茎环区进行划分后识别同源结构域,才能正确处理序列的插入和删除,建立正确的比对以及系统树(Page *et al.*, 2002)。

目前对于线粒体 12S rRNA 二级结构的预测大多数仅局限于结构域 III (Hickson *et al.*, 1996; Page, 2000; Page *et al.*, 2002),而对 16S rRNA 二级结构的预测仅局限于结构域 IV 和 V (Buckley *et al.*, 2000)。这主要是由于这些结构域的序列高度保守,便于设计引物进行 PCR 扩增和测序。昆虫中只有黑果蝇 *Drosophila virilis* Sturt (Cannone *et al.*, 2002)、意大利蜜蜂 *Apis mellifera* (Gillespie *et al.*, 2006)及疑钩额蚤 *Ruspolia dubia* (Zhou *et al.*, 2007)构建了两类 rRNA 完整序列的二级结构。

目前直翅目只有 3 种线粒体基因组全序列的报道(Flook *et al.*, 1995; Kim *et al.*, 2005; Zhou *et al.*, 2007)。云斑车蝗 *Gastrimargus marmoratus* (Thunberg) 属于直翅目(Orthoptera)蝗总科(Acridoidea)蝗科(Acrididae),是广泛分布于我国南北各地的常见蝗虫。刘念等(2006)应用长 PCR 技术成功扩增出了包括云斑车蝗在内的 10 种蝗虫的线粒体基因组。在此基础上,本文采用引物步移法

测定了云斑车蝗线粒体基因组全序列,在对序列进行拼接、注释和详细分析的基础上,根据黑果蝇、意大利蜜蜂和疑钩额蚤的 rRNA 二级结构,对云斑车蝗的两种 rRNA 的二级结构进行了预测并与上述物种进行了比较,为线粒体基因组序列和 rRNA 二级结构在进化和系统发育分析中的应用提供进一步的理论依据。

1 材料与方法

1.1 材料

云斑车蝗标本采自陕西省西安市长安县,浸泡于无水乙醇中,保存于 4℃。

1.2 方法

1.2.1 总 DNA 提取:将蝗虫后足股节肌肉置于 600 μ L 匀浆缓冲液(0.01 mol/L Tris, pH 7.8, 5 mmol/L EDTA, 5% SDS (w/v), 50 ng proteinase K/ μ L)中研磨,65℃水浴消化 3~5 h。用平衡酚(pH 7.6~7.8)提纯两次,再用 CI [氯仿:异戊醇(v/v) = 24:1]提纯一次。最后用 -20℃预冷的 100%乙醇沉淀和 70%乙醇洗涤总 DNA。

1.2.2 PCR 扩增:以所提取的总 DNA 为模板,用 2 对长 PCR 引物扩增线粒体 DNA,引物信息和扩增条件见刘念等(2006)。扩增的 2 个相互重叠的片段 A 和 B 覆盖整个线粒体基因组。长 PCR 反应产物用 1.0%的琼脂糖凝胶电泳检测,使用 λ DNA/Hind III 作为 Marker,在紫外透射仪下观察电泳结果,使用 DNA 凝胶回收试剂盒进行纯化。

以长 PCR 的纯化产物为模板进行 Sub-PCR,扩增出 28 条长度为 600~1 400 bp 的短片段。用 DNA 凝胶回收试剂盒进行纯化。

1.2.3 测序:Sub-PCR 产物进行直接测序或者克隆测序,克隆使用的载体为 pMD18-T。测序试剂盒为 ABI PRISM BigDye Terminator v3.1,直接测序反应的引物与 PCR 引物相同。测序仪为 ABI PRISM™ 3100-Avant Genetic Analyzer。

1.2.4 序列拼接、注释和分析:使用 Staden Package 1.7(Bonfield *et al.*, 1995)进行序列拼接和注释,通过与非洲飞蝗 *Locusta migratoria* 线粒体 DNA 序列进行比对确定蛋白质基因、tRNA 基因和 rRNA 基因的位置。使用 MEGA 3.0 统计线粒体基因组中的基因间隔和重叠,碱基组成,蛋白质基因密码子使用频率。使用 tRNAScan-SE ver. 1.21 (Lowe and Eddy, 1997)对 tRNA 二级结构进行预测。

表 1 云斑车蝗线粒体基因组组成

Table 1 The gene contents of <i>Gastrimargus marmoratus</i> mitochondrial genome							
基因 Genes	编码链 Coding strand	起始位置 Start position	终止位置 End position	长度 Length	间隔* Intergenic length*	起始密码子 Initiation codons	终止密码子 Termination codons
<i>tRNA^{Ile}</i>	J	1	65	65	0		
<i>tRNA^{Gln}</i>	N	63	131	69	- 3		
<i>tRNA^{Met}</i>	J	136	198	69	4		
<i>ND2</i>	J	200	1 234	1 035	1	ATG	TAA
<i>tRNA^{Trp}</i>	J	1 233	1 301	69	- 2		
<i>tRNA^{Cys}</i>	N	1 294	1 357	64	- 8		
<i>tRNA^{Tyr}</i>	N	1 360	1 427	68	2		
<i>CO I</i>	J	1 432	2 973	1542	4	TTG	TAG
<i>tRNA^{Leu}</i>	J	2 969	3 034	66	- 5		
<i>CO II</i>	J	3 043	3 714	672	8	ATG	TAA
<i>tRNA^{Asp}</i>	J	3 713	3 776	64	- 2		
<i>tRNA^{Lys}</i>	J	3 780	3 850	71	3		
<i>ATP8</i>	J	3 868	4 026	159	17	ATC	TAA
<i>ATP6</i>	J	4 020	4 697	678	- 7	ATG	TAA
<i>CO III</i>	J	4 702	5 493	792	4	ATG	TAA
<i>tRNA^{Gly}</i>	J	5 496	5 560	65	2		
<i>ND3</i>	J	5 561	5 914	354	0	ATT	TAG
<i>tRNA^{Ala}</i>	J	5 913	5 977	65	- 2		
<i>tRNA^{Arg}</i>	J	5 981	6 044	64	3		
<i>tRNA^{Asn}</i>	J	6 045	6 110	66	0		
<i>tRNA^{Ser}</i>	J	6 111	6 177	67	0		
<i>tRNA^{Glu}</i>	J	6 178	6 244	67	0		
<i>tRNA^{Phe}</i>	N	6 243	6 307	65	- 2		
<i>ND5</i>	N	6 308	8 033	1 719	0	TTG	T
<i>tRNA^{His}</i>	N	8 043	8 107	65	9		
<i>ND4</i>	N	8 111	9 448	1 338	3	ATG	TAA
<i>ND4L</i>	N	9 442	9 723	282	- 7	ATG	TAA
<i>tRNA^{Thr}</i>	J	9 727	9 790	64	3		
<i>tRNA^{Pro}</i>	N	9 791	9 853	63	0		
<i>ND6</i>	J	9 856	10 377	522	2	ATT	TAA
<i>Cytb</i>	J	10 382	11 518	1 131	3	ATG	TAA
<i>tRNA^{Ser}</i>	J	11 524	11 593	70	5		
<i>ND1</i>	N	11 615	12 550	936	11	ATG	TAA
<i>tRNA^{Leu}</i>	N	12 552	12 619	68	1		
16S rRNA	N	12 613	13 939	1 327	- 7		
<i>tRNA^{Val}</i>	N	13 941	14 012	72	1		
12S rRNA	N	14 013	14 847	835	0		
A + T rich		14 848	15 904	1 057	0		

* 间隔一列中,正数表示基因间隔碱基数,负数表示基因重叠碱基数。In the column intergenic length, the positive number indicates interval base pairs between genes, while the negative number indicates the overlapping base pairs between genes.

表 2 云斑车蝗线粒体基因组核苷酸组成

Table 2 The nucleotide composition of <i>Gastrimargus marmoratus</i> mitochondrial genome						
	A (%)	T (%)	C (%)	G (%)	A + T (%)	n
全基因组 Whole genome	42.8	33.2	14.5	9.48	76.0	15 904
蛋白质编码基因 Protein-coding genes	33.2	41.9	12.7	12.2	75.1	11 135
密码子第一位点 1st codon site	33.3	35.2	12.4	19.1	68.5	3 712
密码子第二位点 2nd codon site	20.3	46.0	19.3	14.3	66.3	3 712
密码子第三位点 3rd codon site	46.0	44.4	6.4	3.2	90.4	3 712
tRNA 基因 tRNA genes	39.2	35.3	10.8	14.7	74.5	1 467
rRNA 基因 rRNA genes	32.1	45.3	8.3	14.3	77.4	2 162
A + T 富集区 A + T-rich region	51.9	32.4	10.4	5.3	84.3	1 057

n: 碱基数 Number of bases in corresponding sequence.

表 3 云斑车蝗线粒体基因组密码子使用情况

Table 3 The codon usage of <i>Gastrimargus marmoratus</i> mitochondrial genome							
密码子 (氨基酸) Codon (aa)	使用次数(相对 同义密码子使用) n (RSCU)	密码子 (氨基酸) Codon (aa)	使用次数(相对同 义密码子使用) n (RSCU)	密码子 (氨基酸) Codon (aa)	使用次数(相对 同义密码子使用) n (RSCU)	密码子 (氨基酸) Codon (aa)	使用次数(相对 同义密码子使用) n (RSCU)
UUU (F)	273.0 (1.66)	UCU (S)	93.0 (2.27)	UAU (Y)	170.0 (1.79)	UGU (C)	40.0 (1.74)
UUC (F)	56.0 (0.34)	UCC (S)	13.0 (0.32)	UAC (Y)	20.0 (0.21)	UGC (C)	6.0 (0.26)
UUA (L)	384.0 (4.17)	UCA (S)	112.0 (2.73)	UAA (*)	0.0 (0.00)	UGA (W)	98.0 (1.85)
UUG (L)	31.0 (0.34)	UCG (S)	3.0 (0.07)	UAG (*)	0.0 (0.00)	UGG (W)	8.0 (0.15)
CUU (L)	63.0 (0.68)	CCU (P)	59.0 (1.74)	CAU (H)	47.0 (1.45)	CGU (R)	18.0 (1.33)
CUC (L)	4.0 (0.04)	CCC (P)	11.0 (0.32)	CAC (H)	18.0 (0.55)	CGC (R)	0.0 (0.00)
CUA (L)	65.0 (0.71)	CCA (P)	66.0 (1.94)	CAA (Q)	63.0 (1.88)	CGA (R)	33.0 (2.44)
CUG (L)	6.0 (0.07)	CCG (P)	0.0 (0.00)	CAG (Q)	4.0 (0.12)	CGG (R)	3.0 (0.22)
AUU (I)	321.0 (1.75)	ACU (T)	74.0 (1.49)	AAU (N)	156.0 (1.70)	AGU (S)	27.0 (0.66)
AUC (I)	45.0 (0.25)	ACC (T)	14.0 (0.28)	AAC (N)	28.0 (0.30)	AGC (S)	2.0 (0.05)
AUA (M)	256.0 (1.82)	ACA (T)	109.0 (2.19)	AAA (K)	84.0 (1.70)	AGA (S)	74.0 (1.80)
AUG (M)	26.0 (0.18)	ACG (T)	2.0 (0.04)	AAG (K)	15.0 (0.30)	AGG (S)	4.0 (0.10)
GUU (V)	93.0 (2.09)	GCU (A)	71.0 (1.75)	GAU (D)	66.0 (1.78)	GGU (G)	77.0 (1.41)
GUC (V)	5.0 (0.11)	GCC (A)	6.0 (0.15)	GAC (D)	8.0 (0.22)	GGC (G)	1.0 (0.02)
GUA (V)	78.0 (1.75)	GCA (A)	84.0 (2.07)	GAA (E)	69.0 (1.86)	GGA (G)	131.0 (2.39)
GUG (V)	2.0 (0.04)	GCG (A)	1.0 (0.02)	GAG (E)	5.0 (0.14)	GGG (G)	10.0 (0.18)

n：该密码子的使用次数 Frequency of codon used；RSCU：相对同义密码子使用 Relative synonymous codon usage；*：终止密码子 Stop codon.

2.5 rRNA 二级结构

rRNA 基因的位置在昆虫中比较保守。12S rRNA 基因位于 A + T 富集区的 5'端和 tRNA^{Val}基因的 3'端，16S rRNA 基因位于 tRNA^{Val}基因的 5'端和 tRNA^{Leu}(CUN)基因的 3'端,长度分别为 835 bp 和 1 327 bps参照已经建立的意大利蜜蜂(Gillespie *et al.*，2006)和黑果蝇(Cannone *et al.*，2002) 12S 和 16S rRNA 二级结构,预测并绘制了云斑车蝗 12S 和 16S rRNA 二级结构(图 2 和图 4)。12S rRNA 二级结构由三个结构域共 30 个茎环组成。结构域Ⅰ与结构域Ⅱ在所比较的几个物种中均有变化,结构域Ⅲ则比较保守(图 2 和图 3)。16S rRNA 二级结构由 6 个结构域共 44 个茎环组成,与其他物种相比,仅在某些位置茎的长度与环的大小略有差异(图 4)。

3 讨论

3.1 蛋白质基因的起始和终止

COⅠ基因的起始密码子在节肢动物的线粒体基因组中具有明显的多样性,有 3 种四联体起始密码子 ATAA,TTAA 和 ATTA(Flook *et al.*，1995；Lewis *et al.*，2005)；双翅目昆虫中还发现与 TCG 起始密码子相邻的六联体起始信号 ATTTAA(Junqueira *et al.*，2004)。本研究 COⅠ没有发现非洲飞蝗那样的四联体,而是以 TTG 作为其起始密码。后者曾经

在四斑按蚊(Mitchell *et al.*，1993)，衣鱼 *Tricholepidion gertschi*(Nardi *et al.*，2003)，萤火虫 *Pyrocoelia rufa*(Bae *et al.*，2004)和几种线虫中发现(Okimoto *et al.*，1992)。此外,ND5 基因也使用了 TTG 作为起始密码子。推测蛋白质基因中的不正常起始密码子在转录成 mRNA 后会通过 RNA 编辑转变为正常起始密码子(Kadowaki *et al.*，1995),但这一机制需要通过测定 COⅠ基因的 cDNA 序列来检验。

ND5 基因使用了不完全终止密码子 T。非洲飞蝗中有 4 个基因为不完全终止密码子,东方蝼蛄有 10 个,更为频繁。不完全终止密码子(T 或 TA)在无脊椎动物甚至是哺乳动物线粒体 DNA 中都很常见(Bibb *et al.*，1981；Clary and Wolstenholme,1985)，对此的解释是转录的 mRNA 会动用其 PolyA 尾完成终止密码子从而终止翻译(Anderson *et al.*，1981)。

3.2 tRNA 二级结构

tRNA^{Ser}(AGN)缺少 DHU 臂,只有由 11 个碱基组成的环。这种情况在包括昆虫在内的其他后生动物中比较常见(Wolstenholme,1992)。尚不清楚成熟的 tRNA^{Ser}(AGN)是否具有相同的特征,以及这种异常对于其转运功能造成的影响,或者有可能在 tRNA^{Ser}(UCN)存在的情况下,该产物根本没有活性,这些有待于通过实验来解释。

在臂和环的长度上,DHU 臂和 T ψ C 臂在所有

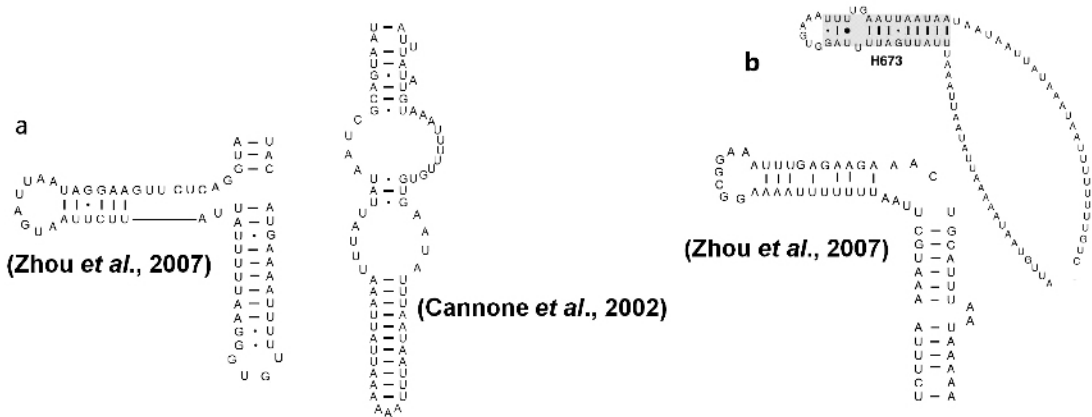
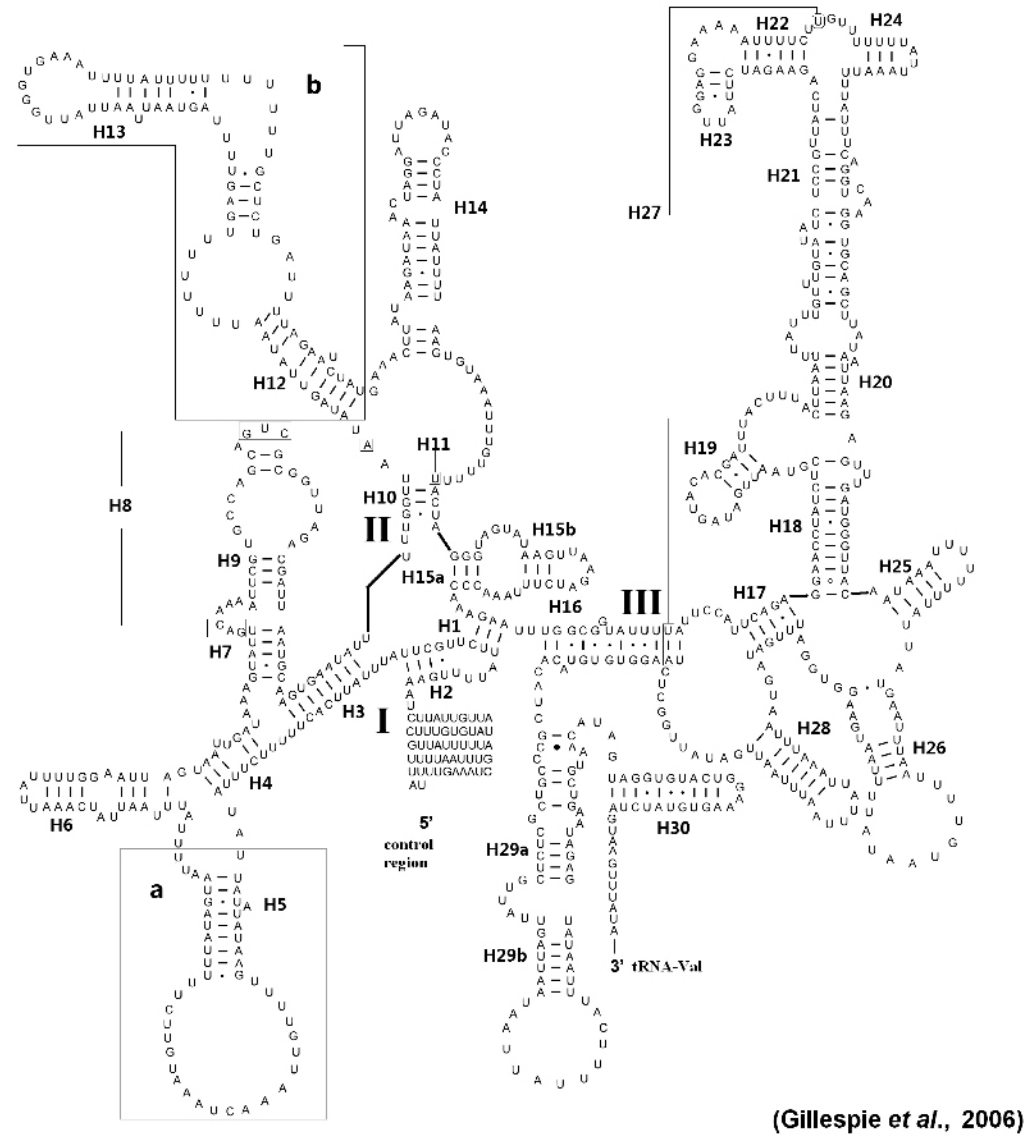


图 2 云斑车蝗线粒体 12S rRNA 二级结构

Fig. 2 The 12S rRNA secondary structure of *Gastrimargus marmoratus* mitochondrial genome
方框标出的 a,b 两处表示在所参考的几个种中的可变部分。
The a and b region in panes mark the variable sections among the species compared.

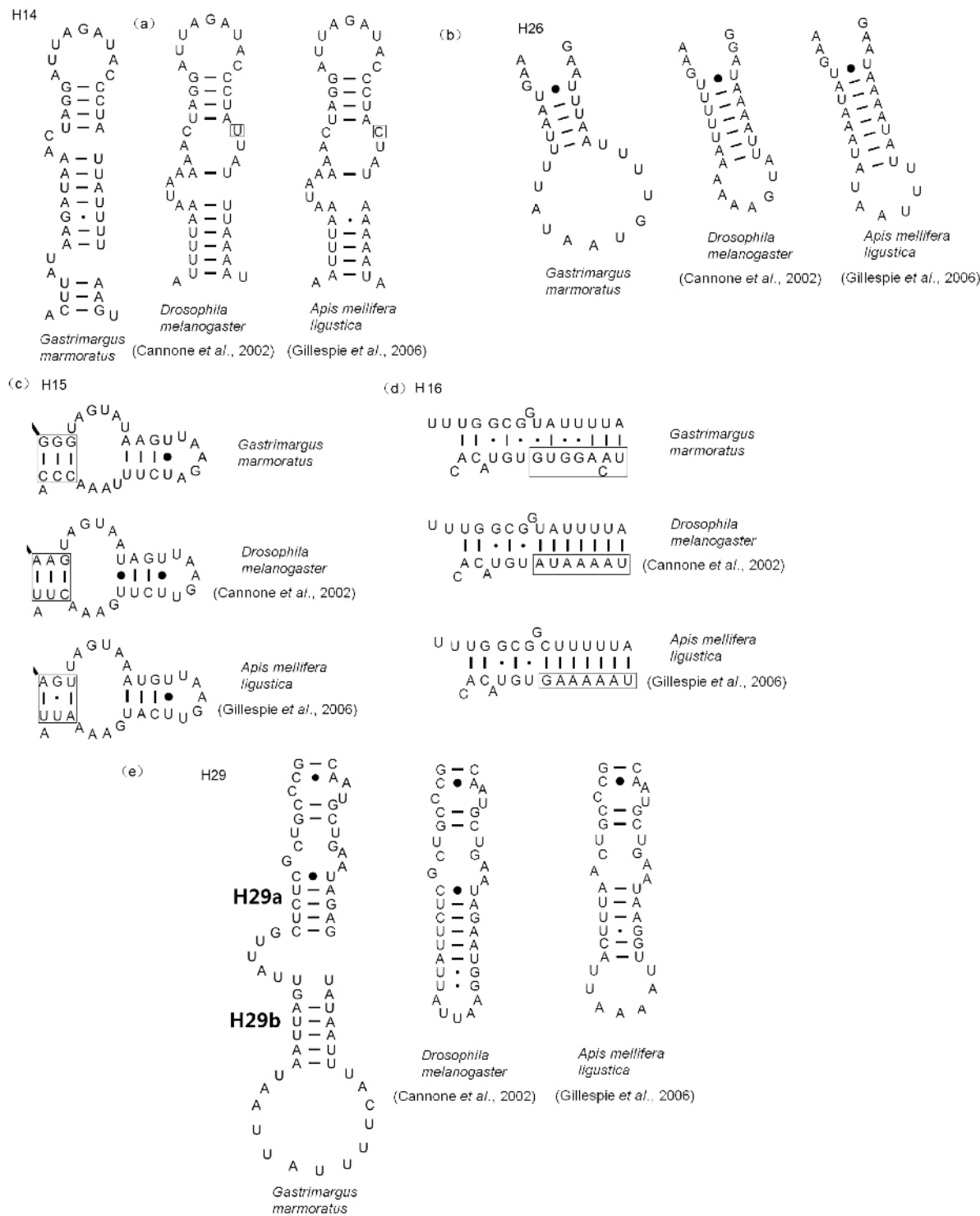


图 3 云斑车蝗、黑果蝇和意大利蜜蜂线粒体 12S rRNA 部分区段的比较

Fig. 3 The comparison of 12S rRNA secondary structures among *Gastrimargus marmoratus*, *Drosophila virilis* and *Apis mellifera*

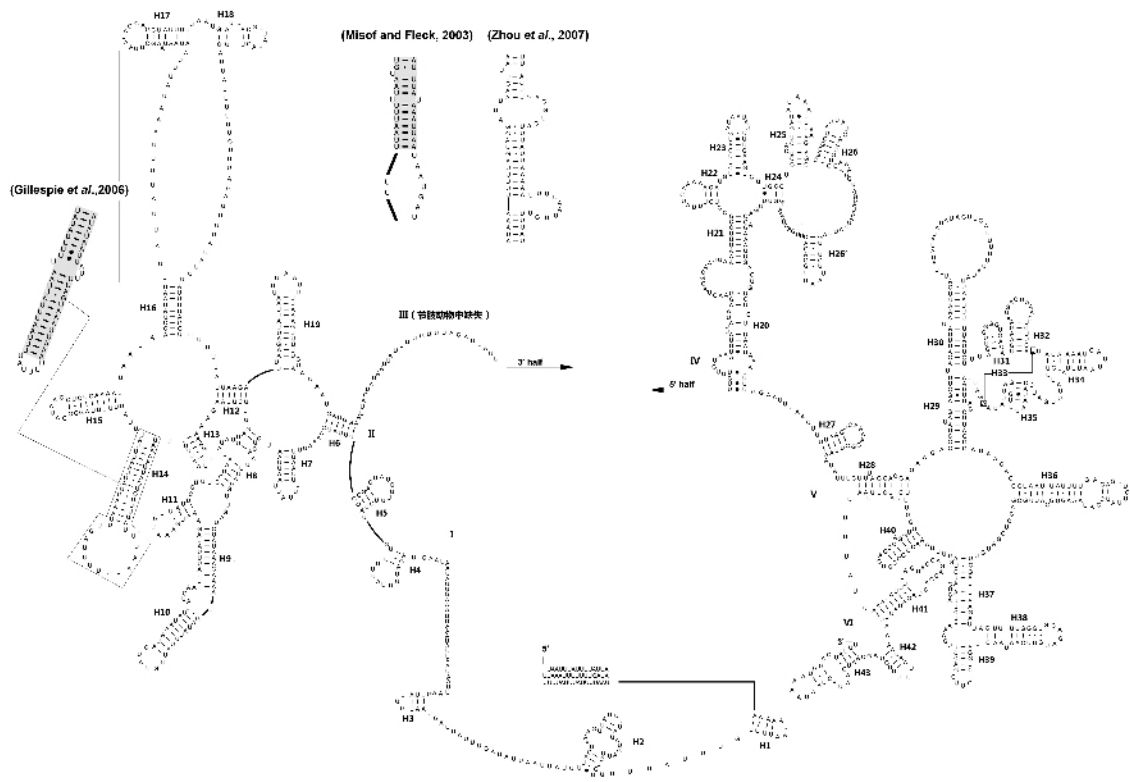


图 4 云斑车蝗 16S rRNA 二级结构

Fig. 4 The secondary structure of 16S rRNA in *Gastrimargus marmoratus* mitochondrial genome
方框与虚线表示该区域在几个种中的不同之处 The pane and dashed indicate the variable sections among the species referred.

tRNA 中变化较大,氨基酸接受臂和反密码子臂则相对保守,分别为稳定的 7 对和 5 对碱基。但也有例外。在所预测到的 tRNA^{Ser}(UGA)二级结构中,氨基酸接受臂 3'端有一个碱基 U 不能配对,形成凸环(bulge);另有一处异常情况是其反密码子环上为 9 个碱基,而不是通常的 7 个碱基。非洲飞蝗和东方蝼蛄中没有发现这样的异常,但在最近研究的蜘蛛 *Pseudocellus pearsei* 线粒体基因组的 tRNA^{Ala} 和 tRNA^{Phe} 中都有出现(Fahreïn et al., 2007)。

云斑车蝗 tRNA 二级结构中的 19 对碱基错配介于非洲飞蝗(4 对)和东方蝼蛄(34 对)之间。频率最高的为 GU 配对(14)。有人认为线粒体基因组 tRNA 基因的部分错配可以通过 RNA 编辑校正,不会引起氨基酸转运上的障碍(Yokobori and Pääbo, 1995)。

tRNA^{Ser}(AGN)缺少 DHU 臂,只有由 11 个碱基组成的环。这种情况在包括昆虫在内的其他后生动物中比较常见(Wolstenholme, 1992)。tRNA^{Ser}(AGN) DHU 臂的长度由低等无脊椎动物到高等脊椎动物逐渐变短,对应的是密码子使用频率降低。对哺乳动物这类 tRNA 的结构研究表明,它们可以形成潜在的 L 型结构以维持反密码子与 CCA 接受臂之间

的距离(Hanada et al., 2000)。尚不清楚成熟的 tRNA^{Ser}(AGN)是否具有相同的特征,以及这种异常对于其转运功能造成的影响。

3.3 A + T 富集区

A + T 富集区是后生动物线粒体基因组中的一段非编码区,由于其富含 A + T,称为 A + T 富集区(Zhang and Hewitt, 1997)。脊椎动物 A + T 富集区包含有线粒体 DNA 重链复制的起点(Brown, 1985),各种果蝇中则含有双链复制的起点(Fauron and Wolstenholme, 1980)。A + T 富集区在脊椎动物和无脊椎动物中线粒体基因组中的位置比较保守,位于 12S rRNA 基因和 tRNA-Ile 基因之间,但它的长度在各个物种中却变化很大,在昆虫中最短 70 bp(Zhou et al., 2007),最长可达 4 061 bp (Lewis et al., 1995)。在近缘种类之间 A + T 富集区的长度差异主要是由于串联重复拷贝的数目不同,虽然也有插入和缺失的存在,但影响相对较小(Zhang and Hewitt, 1997)。

在云斑车蝗 A + T 富集区发现了 3 个串联重复序列:Rpt1, Rpt2 和 Rpt3。在非洲飞蝗线粒体基因组中也发现了这样的情况(Flook et al., 1995),但只

有 2 个重复单位。云斑车蝗的 Rpt1 的 5' 端起始于 14 791, 伸入 12S rRNA 基因的 3' 端(14 847), 有 54 bp 的重叠, 和非洲飞蝗相似。Rpt2 和 Rpt1 长度均为 166 bp, 核苷酸序列完全相同。Rpt3 长 120 bp, 与 Rpt1 和 Rpt2 的前 120 bp 相比只有一个碱基不同。和非洲飞蝗一样, 3 个重复单位之间没有间隔。重复序列的 A + T 含量为 80.1%, 略低于整个 A + T 富集区。据推测, 这种重复单位数量的异质性可能是由于复制过程中的滑移事件所致(Zhang and Hewitt, 1997)。

3.4 rRNA 二级结构

12S rRNA 二级结构显示结构域 I 与结构域 II 在所比较的几个物种中均有变化, 结构域 III 则比较保守。结构域 I 中的 H5 与蜜蜂的对应区域 H47 相似, 而在疑钩额蚤(Zhou *et al.*, 2007) 与黑果蝇(Cannone *et al.*, 2002) 中这一区段均有变化(图 2 : a)。结构域 II 中的 H12 则与黑果蝇(Cannone *et al.*, 2002) 对应区段类似, 在意大利蜜蜂中该位置仅形成一个长度仅 4 bp 且有错配的茎(Gillespie *et al.*, 2006), 疑钩额蚤的这一区域 H12' 的茎区最长, 更说明了这一区域的可变程度(图 2 : b)。H14 与蜜蜂和果蝇相比, 茎的形状有一些变化而环区在 3 个物种中完全相同(图 3 : a); 多数茎区的碱基补偿性突变可以保证二级结构的保守性不受一级序列变化的干扰, 如 H15(图 3 : c), 而在 H16 中, 配对区的两部分相对于果蝇和蜜蜂并没有发生共变(covariation), 只有下游发生了变化, 但仍然可以形成非常规配对(图 3 : d)。H26 不能形成像果蝇和蜜蜂那样长的茎区(图 3 : b), H29 的茎区有较长的插入, 环区也比果蝇和蜜蜂大得多(图 3 : e)。

16S rRNA 大体骨架与意大利蜜蜂的模板相似, 仅在某些位置茎的长度与环的大小略有差异(图 4)。结构域 I 中的 5 个茎环(H1, H2, H3, H4, H5) 都较短而且彼此距离较远, 结构域 II 中 H14 的茎的长度比意大利蜜蜂对应位置 H837 短(图 4); 方框所圈定的区域在蜻蜓目昆虫中(Misof and Fleck, 2003) 和疑钩额蚤中(Zhou *et al.*, 2007) 都可以形成茎环。结构域 III 在节肢动物中不存在, 仅有一段未形成二级结构的 30 bp 左右的序列。结构域 IV 和结构域 V 与意大利蜜蜂较一致。

通过预测 12S 和 16S rRNA 的二级结构信息, 有助于了解这两个基因不同区段的功能以及对应的不同进化速率, 从而根据需求选择合适的基因区段来研究特定分类阶元的系统发育问题。

致谢 感谢西北工业大学生命科学学院卢慧莹博士为实验设计通用引物。

参 考 文 献 (References)

- Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG, 1981. Sequence and organization of the human mitochondrial genome. *Nature*, 290 : 457 – 465.
- Bae JS, Kim I, Sohn HD, Jin BR, 2004. The mitochondrial genome of the firefly, *Pyrocoelia rufa* : complete DNA sequence, genome organization, and phylogenetic analysis with other insects. *Molecular Phylogenetics and Evolution*, 32 : 978 – 985.
- Ballard JW, Olsen GJ, Faith DP, Odgers WA, Rowell DM, Atkinson PW, 1992. Evidence from 12S ribosomal RNA sequences that onychophorans are modified arthropods. *Science*, 258 : 1 345 – 1 348.
- Bibb MJ, Van Etten RA, Wright CT, Walberg MW, Clayton DA, 1981. Sequence and gene organization of mouse mitochondrial DNA. *Cell*, 26 : 167 – 180.
- Black WC IV, Piesman J, 1994. Phylogeny of hard- and soft-tick taxa (Acari : Ixodida) based on mitochondrial 16S rDNA sequences. *Proc. Natl. Acad. Sci. USA*, 91 : 10 034 – 10 038.
- Bonfield JK, Smith KF, Staden R, 1995. A new DNA sequence assembly program. *Nucleic Acids Res.*, 24 : 4 992 – 4 999
- Buckley TR, Simon C, Flook PK, Misof B, 2000. Secondary structure and conserved motifs of the frequently sequenced domains IV and V of the insect mitochondrial large subunit rRNA gene. *Insect Mol. Biol.*, 9 : 565 – 580.
- Cannone JJ, Subramanian S, Schnare MN, Collett JR, D Souza LM, Du Y, Feng B, Lin N, Madabusi LV, Muller KM, Pande N, Shang Z, Yu N, Gutell RR, 2002. The comparative RNA web (CRW) site : an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics*, 3 : 2.
- Clary DO, Wolstenholme DR, 1985. The mitochondrial DNA molecule of *Drosophila yakuba* : nucleotide sequence, gene organization, and genetic code. *J. Mol. Evol.*, 22 : 252 – 271.
- Fahren K, Talarico G, Braband A, Podsiadlowski L, 2007. The complete mitochondrial genome of *Pseudocellus pearsei* (Chelicerata : Ricinulei) and a comparison of mitochondrial gene rearrangements in Arachnida. *BMC Genomics*, 8(1) : 386.
- Flook PK, Rowell CH, Gellissen G, 1995. The sequence, organization, and evolution of the *Locusta migratoria* mitochondrial genome. *J. Mol. Evol.* 41(6) : 928 – 941.
- Gillespie JJ, Johnston JS, Cannone JJ, Gutell RR, 2006. Characteristics of the nuclear (18S, 5.8S, 28S and 5S) and mitochondrial (12S and 16S) rRNA genes of *Apis mellifera* (Insecta : Hymenoptera) : structure, organization, and retrotransposable elements. *Insect Mol. Biol.*, 15 : 657 – 686.
- Hanada T, Suzuki T, Watanabe K, 2000. Translation activity of mitochondrial tRNA with unusual secondary structure. *Nucleic Acids Symposium Series*, 44 : 249 – 250.
- Hickson RE, Simon C, Cooper A, Spicer GS, Sullivan J, Penny D, 1996.

- Conserved sequence motifs, alignment, and secondary structure for the third domain of animal 12S rRNA. *Mol. Biol. Evol.*, 13 :150 – 169.
- Junqueira ACM, Lessinger AC, Torres TT, da Silva FR, Vettore AL, Arruda P, Azeredo Espin AM, 2004. The mitochondrial genome of the blowfly *Chrysomya chloropyga* (Diptera : Calliphoridae). *Gene*, 339 :7 – 15.
- Kadowaki K, Ozawa K, Kazama S, Kubo N, Akihama T, 1995. Creation of an initiation codon by RNA editing in the coxI transcript from tomato mitochondria. *Curr. Genet.*, 28(5) :415 – 422.
- Kambhampati S, Kjer KM, Thorne BL, 1996. Phylogenetic relationship among termite families based on DNA sequence of mitochondrial 16S ribosomal RNA gene. *Insect Mol. Biol.*, 5 :229 – 238.
- Kim I, Cha SY, Yoon MH, Hwang JS, Lee SM, Sohn HD, Jin BR, 2005. The complete nucleotide sequence and gene organization of the mitochondrial genome of the oriental mole cricket, *Gryllotalpa orientalis* (Orthoptera : Gryllotalpidae). *Gene*, 353 :155 – 168.
- Lewis DL, Farr CL, Kaguni LS, 1995. *Drosophila melanogaster* mitochondrial DNA, completion of the nucleotide sequence and evolutionary comparisons. *Insect Mol. Biol.*, 4 :263 – 278.
- Liu N, Hu J, Huang Y, 2006. Amplification of grasshoppers complete mitochondrial genomes using long PCR. *Chinese Journal of Zoology*, 41 (2) :61 – 65. [刘念, 胡婧, 黄原, 2006. 应用长 PCR 扩增蝗虫线粒体全基因组. *动物学杂志*, 41 (2) :61 – 65]
- Lowe TM, Eddy SR, 1997. tRNAscan-SE : a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.*, 25 :955 – 964.
- Misof B, Fleck G, 2003. Comparative analysis of mt LSU rRNA secondary structures of odonates : structural variability and phylogenetic signal. *Insect Mol. Biol.*, 12(6) :535 – 547.
- Mitchell SE, Cockburn AF, Seawright JA, 1993. The mitochondrial genome of *Anopheles quadrimaculatus* species A : complete nucleotide sequence and gene organization. *Genome*, 36 :1 058 – 1 073.
- Nardi F, Spinsanti G, Boore JL, Carapelli A, Dallai R, Frati F, 2003. Hexapod origins : monophyletic or paraphyletic ? *Science*, 299 :1 887 – 1 889.
- Navajas M, Gutierrez J, Lagnel J, Boursot P, 1996. Mitochondrial cytochrome oxidase I in tetranychid mites : a comparison between molecular phylogeny and changes of morphological and life history traits. *Bull. Entomol. Res.*, 86 :407 – 417.
- Okimoto R, Macfarlane JL, Clary DO, Wolstenholme DR, 1992. The mitochondrial genomes of two nematodes, *Caenorhabditis elegans* and *Ascaris suum*. *Genetics*, 130 :471 – 498.
- Page RDM, 2000. Comparative analysis of secondary structure of insect mitochondrial small subunit ribosomal RNA using maximum weighted matching. *Nucleic Acids Res.*, 28 :3 839 – 3 845.
- Page RDM, Cruickshank R, Johnson KP, 2002. Louse (Insecta : Phthiraptera) mitochondrial 12S rRNA secondary structure is highly variable. *Insect Mol. Biol.*, 11 :361 – 369.
- Wolstenholme DR, 1992. Animal mitochondrial DNA : structure and evolution. *Int. Rev. Cytol.*, 141 :173 – 216.
- Yokobori S, Pääbo S, 1995. Transfer RNA editing in land snail mitochondria. *Proc. Natl. Acad. Sci. USA*, 92 :10 432 – 10 435.
- Yukuhiro K, Sezutsu H, Itoh M, Shimizu K, Banno Y, 2002. Significant levels of sequence divergence and gene rearrangements have occurred between the mitochondrial genomes of the wild mulberry silkmoth, *Bombyx mandarina*, and its close relative, the domesticated silkmoth, *Bombyx mori*. *Mol. Biol. Evol.*, 19 :1 385 – 1 389.
- Zhang DX, Hewitt GM, 1997. Insect mitochondrial control region : a review of its structure, evolution and usefulness in evolutionary studies. *Biochem. Syst. Ecol.*, 25 :99 – 120.
- Zhou Z, Huang Y, Shi F, 2007. The mitochondrial genome of *Ruspolia dubia* (Orthoptera : Conocephalidae) contains a short A + T-rich region of 70 bp in length. *Genome*, 50(9) :855 – 866.

(责任编辑 : 赵利辉)